

**Machine translation JP2000516006****DETAILED DESCRIPTION****(19)Publication country**Japan Patent Office (JP)**(12)Kind of official gazette**Published Japanese translations of PCT international publication for patent applications (A)**(11)Official announcement number \*\***table 2000-516006 (P2000-516006A)**(43)Announcement date**November 28, Heisei 12 (2000.11.28)**(54)Title of the Invention**A method and a device to which it shows video contents by displaying two or more key-frames in parallel**(51)The 7th edition of International Patent Classification**

G06F 17/30

H04N 5/76

5/93

**FI**

G06F 15/40 370 D

H04N 5/76 B

G06F 15/403 380 F

H04N 5/93 Z

**Request for Examination**Unrequested**Preliminary examination claim** Unrequested**Number of Pages**24**(21)Application number**Japanese Patent Application No. 10-529208**(86)(22)Filing date**January 22, Heisei 10 (1998.1.22)**(85)Translation presentation day** October 2, Heisei 10 (1998.10.2)**(86)International application number** PCT/IB98/00091**(87)International publication number** WO98/34182**(87)International publication day** August 6, Heisei 10 (1998.8.6)**(31)Application number of the priority**97200286.9**(32)Priority date**February 3, Heisei 9 (1997.2.3)**(33)Country Declaring Priority**European Patent Office (EP)**(31)Application number of the priority**97201802.2**(32)Priority date**June 16, Heisei 9 (1997.6.16)**(33)Country Declaring Priority**European Patent Office (EP)**(81)Designated state** EP(AT,BE,CH,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE),JP,KR**(71)Applicant****Name**KONIN Klecka Philips Electronics N VI**Address**Netherlands 5621 Baer AIN dauphin FURUNEVAUTSUWEHHA 1**(72)Inventor(s)****Name**SEKKAREKI Marco Polo**Address**Netherlands 5656 AA AIN dauphin Prof Horst Ran 6**(74)Attorney****Patent Attorney****Name**Akihide Sugimura (besides five persons)

**(57)Abstract**

While guiding video contents by displaying two or more key-frames of one or more rows in parallel, it is the method of controlling, while mapping of the key-frame which accessed actual access to said video contents in this way by alternative access of the displayed key-frame is expressed. The 1st operational mode that arranges a key-frame in the user interface organization with this single method in the order which aligned in time on the screen, It has the feature which can choose between the continuation key-frames displayed between the 2nd operational mode that arranges the key-frame which has much selectable granularity.

**Claim(s)**

It is a proposal about video contents by displaying two or more key-frames of 1.1 or more rows characterized by comprising the following in parallel. While carrying out inner, It is said video by alternative access of a displayed key-frame. Mapping of a key-frame which accessed actual access to the contents in this way How to control while expressing

**in single user interface organization**, it is a screen about a key-frame. The 1st operational mode arranged in an order which aligned in time upwards

A continuation key displayed. Selectable granularity uneven to inter-frame

2. To key-frame located at center in time in said turn which aligned in time Claim 1 reproducing a related audio interval one by one A method of a statement.
3. a sequence of a key-frame in which a sequential audio interval carried out discrete separation audio expression which received and continued mostly is constituted -- being according to claim 2 - - A method.
4. O relevant to key-frame actually accessed in the 2nd operational mode A method of reproducing - DIO interval according to claim 1.
5. At the same time it carries out highlighting of the key-frame chosen now, This key Many frames are expanded to a format of a double size to other key-frames, furthermore -- detecting the harmful video INTA racing effect -- such an effect -- detection -- Make such an effect reduce by vertical declination in a \*\*\*\* case. A method according to claim 1 by which it is characterized.
6. At the same time it carries out highlighting of the key-frame chosen now, this key many frames are expanded to a format of a double size to other key-frames -- performing rise sampling filter processing to a picture further, before displaying a picture A method according to claim 1 by which it is characterized.
7. Method of displaying subtitle or other pertinent information which were extracted to related sequence of key-frame or key-frame according to claim 1.
8. Device constituted so that method indicated to claim 1 might be performed.

**Detailed Description of the Invention**

In the method and background this invention of a device invention to which it shows video contents, two or more key-frames of one or more rows are displayed in parallel by displaying two or more key-frames in parallel.

Therefore, while guiding video contents, it is related with the method of controlling, while mapping of the key-frame which accessed actual access to said video contents in this way by alternative access of the displayed key-frame is expressed.

Using the typical portion of the video presentation recorded for next selective reproduction as a key-frame is proposed. The continuation video stream can mean that video continues being "one", and animation, a series of still pictures, or the interactive sequence of a picture can be included in this stream. The character can be used as various things, for example, a movie, news, or a shopping list.

This art is indicated by S.W.Smolair and paper "Content-Based Video Indexing and Retrieval" of H.J.Zhang, IEEE Multimedia, Summer 1994, and pp.62-72.

A key-frame can be taken out from this material with a derivation algorithm in a user's house at the time of reception of video material, or for example, by a video provider, it can attach a label to a key-frame so that each video shot may begin by a key-frame. The 3rd method makes these frames continue mutually by the uniform time interval relevant to standard video speed. This invention carries out key-frame use, gives a user the dynamic OBA view of a video presentation with useful facility, and is based on recognition that it is necessary to make further access to video material, selection of the next display, selection release, or edit into \*\*.

Generally the memory to a mass medium does not permit immediate access, but the specific problem over the present project over digital one of a video image and compression encoding has it in the point that the linear recording density which can be especially expressed with the frame number per unit memory size is uneven. Adding the secondary memory medium which has a high access speed although it is small capacity to a high capacity main memory medium like a tape is proposed. In this case, the processing which edits video material into a summarized type, a changed type, or a rearrangement form at the next high-speed passing <a thing> on, execution of a trick mode like high-speed backward feed , and display sake sees from the both sides of the point of a user interface, and the point of storing technology, and produces a remarkable difficulty. There are an outline of an invention, therefore the purpose of this invention in giving a user the much more natural feeling of memory organization and video material while introducing still higher flexibility into user interface organization especially. in the user interface organization with the single feature of this invention for this purpose, It is in the ability to choose between the 1st operational mode that arranges a key-frame in the order which aligned in time on the screen, and the 2nd operational mode that arranges the key-frame which has two or more uneven selectable granularity between the continuation key-frames displayed. If a key-frame is shown in the order which aligned in time, when these frames will continue mutually by the uniform time interval relevant to standard video speed, for example, high-speed passing <a thing> on and high-speed backward feed can be performed easily. An easy, easy change of the layer level which has the variable granularity in inter-frame can perform easy, easy access and edit. A key-frame, when taking out the part from the start part of a movie shot at least, or also when taking out from the related event of the others generated by the original film editor, the same thing can say. Thus, clustering processing can be performed automatically.

While carrying out highlighting of this frame by expanding many key-frames chosen now in a double format to other key-frames in this invention method, The harmful video INTARESU effect is detected, if that is right, this effect will be reduced by vertical decimation, and rise sampling filter processing is performed to a picture before a display again. Although it turns out on experience that it can approve, the video distortion of a comparatively small key-frame needs to lecture on an additional means for a picture improvement, when expanding a specific key-frame. this invention person confirmed that it brought about the comfortable and useful improvement in image quality although this upgrading does not necessarily reach the image quality under a normal condition.

This invention relates also to the device constituted so that the method mentioned above might be enforced. Other features of this invention are indicated to the dependent claim.

These features, other features, and advantage of a brief explanation of the drawings this invention are explained in detail below about a preferred embodiment with reference to drawings. In a drawing, drawing 1 is a block diagram of TV-recorder combination device, Drawing 2 shows the typical structure of video record, and drawing 3 shows the example of a design of a scroll mosaic user interface, Drawing 4 shows the example of a design of a scroll list user interface, drawing 5 shows the example of a design of a more extensive graphical user interface, drawing 6 shows the display of a subtitle, and drawing 7 shows the constitutional diagram of a system action.

Although it is related to detailed explanation, especially the usual customer of a preferred embodiment, and use of a private home, various advantages which are not what is limited to the thing related to such use are as follows.

- A key-frame must be made to show so that the user located in a typical TV observation distance can identify these mutually.
- The number of the key-frames shown simultaneously should presuppose that it is enough to give a user the Oba view of the significant portion of the contents of the Digital Video material.

- The key-frame should be displayed in the undistorted form by holding an aspect ratio, for example.
- It is preferred to make it the remote control device of a TV set itself operate as a user control apparatus.
- Feedback information must be able to be recognized from typical observation distance.
- Generally make unnecessary a computer concept like a "drug" and "drops." - Various functions are not continuous, and they must be realized so that it may become **being used occasionally and**.

- The user interface should carry out the table of the linear model good **video** and known.

The explanatory view 1 of a specific example is a lineblock diagram showing TV-recorder combination device by this invention. The element 20 shows the immediate-control and power supply section which TV-set-displays and is related. The element 22 shows the terminal area which has an antenna or a signal distribution entity type **other** like cable distribution. The entity which takes out the Digital Video information or a digital signal portion from an input signal can also be suitably included in this element. The element 34 shows the video stream of various intersubsystems of drawing 1, and the related routing part of information. Routing is controllable by the control box 28 with a control signal via the line 35. Although the line 35 is shown as a single bidirectional interconnection line, it shall consist of arbitrary numbers of one ways or bidirectional lines actually. The control box 28 receives a detecting signal from other subsystems 38 and 40 while receiving a detecting signal through the line 30 from the display 20, and it controls these subsystems. The block 38 is a linear tape recorder which has a very high storage capacity of a multi-G byte field. Although the block 40 is a magnetic disk recorder which has a high storage capacity and this recorder only has a part of storage capacity of the recorder 38, the access speed of this recorder is farther **than the recorder 38** quick by crossing track jump access. The block 38 and the block 40 constitute 2 level memory organization similar to a computer memory cache system together, and all the items of a video presentation are memorized once **at least**. The element 24 shows the remote control device which communicates with the subsystem 28 and other subsystems 38 and 40 indirectly while communicating via the display device 20 and the wireless path 26.

Drawing 2 shows the typical composition of a video presentation. In order to use video contents as a functional target, the bar 60 is a form of a frame or contains the video itself as a string of compression video contents like MPEG numerals. Although information is memorized with advance of video time along with a bar, it is not necessary to make the actual requirements for memory uniform **regeneration time**. The key-frame with which it is dotted is shown by black vertical stripe like 68. Each key-frame is used as the thing showing all the videos in the interval to the following key-frame, or a thing representing this. as for a key-frame, a video provider attaches a label -- or" -- table OBU contents -- " (TOC) -- it can choose as the 1st frame of each new shot by inserting. or -- again -- a receiver -- a certain algorithm -- a video content -- FUREMUHE of one frame to the next -- changing rapidly is detectable. This algorithm assumes this invention to be a right thing. As shown in a figure, distribution of a key-frame can be made uneven. Other mechanisms are in the interval of regulation of a continuous key-frame, for example, make it continue mutually every 2 to 3 seconds. This example shows only a key-frame to the indicator 62. A key-frame is composed a little by the layered structure, and the indicator 64 shows only the key-frame of a high level of the limited lot.

This layered structure shall be used as many levels, and the indicator 66 shall be connected only with a single the whole video presentation 60 key-frame. Various levels of a key-frame can be defined within the organization mentioned above and different organization, and can also be put in order.

While memorizing the main part of a video presentation on the tape recorder 38 in drawing 1, memory mapping, It can carry out so that it may play with the disk recorder 40 together with the short video just behind a key-frame related if a key-frame is made at least, and/or an audio interval. The length of such an interval can be made to be able to respond to the waiting time of the linear tape recorder 38, and can attain real-time access. A video presentation can be intrinsically made linear like a movie. In other use, other pictures used for a predetermined memory interval by animation, a still picture, or the consumer can be included.

A certain key-frame can oppress this. This combines effectively the time interval in front of the

key-frame concerned with a subsequent time interval. An interval is separable again by a reset function. The key-frame of various classes, for example, the key-frame of the class mutually separated by the fixed time interval, can be oppressed. The key-frame introduced by the provider to the key-frame of various different classes, for example, the key-frame generated by a local algorithm at the time of reception, can be used for one presentation.

Drawing 3 shows the example of a design of a scroll mosaic user interface. Every screen departs from the upper left and presents 20 key-frames to the lower right. Each key-frame has a number of the comprehensive ranking of the key-frame of a graphic display. Actually, highlighting of the key-frame 144 is carried out by rectangular control cursor. The user can drive remote control and can move freely the button top displayed on the bar displayed on the key-frame top which had cursor displayed by the navigation control section of a cursor device, the upper part, and a pars basilaris ossis occipitalis. If a user moves control cursor to the left in an upper left corner, a display will jump only 20 key-frames back. If it is made to move to the right in a lower right corner, a display will jump only 20 frames in front. Access of other portions of the presentation divided into the portion of five equal length by accessing the upper bar of a screen is controllable. A black horizontal bar shows the full time covered by 20 key-frames displayed on here of all the presentations.

Other functions can be made to start by choosing a specific key-frame first and subsequently choosing 1 \*\* of a pars-basilaris-ossis-occipitalis button. A "view program" (program view) button controls the start by the key-frame accessed with cursor. Although the "View segment" (segment view) button is the same as that of a view program, only the unsegment ended by the following key-frame is reproduced. A "view from x to y" (view from x to y) button controls a start with the frame of the earlier one in time of the two key-frames accessed with cursor, and a stop with the frame of the later one in time. Other modes are realizable with a key-frame function preselection capability. For example, Fast-forward (high-speed passing <a thing> on) or slow-forward (low-speed passing <a thing> on) which enables it to check generating of an interval with a specific user. Or fast/slow reverse (a high speed / low-speed backward feed) which attains the predetermined video effect can be attained. While on display, at the time of the momentary passage relevant to a specific key-frame, a video stream is effectively displayed until this key-frame becomes active and reaches the instant relevant to the following key-frame. At the time of attainment of the following key-frame, this frame turns into an active frame. With such a function, a user a VCR by deleting a predetermined segment like commercials as opposed to an interval display sequence. It becomes possible to program to a straight forward so that cautions turn to a predetermined detail according to low-speed passing <a thing> on. Activate an audio with the control button which is not illustrated while on display, it oppresses, or things are made smoothly. Or although it continues controlling an audio, video cursor can be made discrete and it can only be supposed again that it is to carry out a step to an interval from an interval with suitable highlight directions.

Drawing 4 shows the typical example of a design of a scroll list user interface. In this example, all the screens have five key-frames displayed on the pars basilaris ossis occipitalis, and highlighting is carried out by the control cursor of the rectangle which the key-frame 145 moves along with the edge. The key-frame 145 is expressed as big magnification also in a background. The position of a button is different although this control interface is the same as that of the thing of drawing 3. An expansion key-frame can also be oppressed in a multi key frame bar.

Drawing 5 shows a more extensive graphical user interface. A column on either side play (reproduction), stop (stop), select (selection), **1st** It is a column of the control button for cut (cutoff), paste (stick), fast reverse (high-speed backward feed), zoom+ (zoom +), zoom- (zoom), and fast forward (high-speed passing <a thing> on). The line of a pars basilaris ossis occipitalis has a sequence of nine key-frames relevant to the scene different, respectively or shot which hardly has correlation. By the step moving within the layered structure of a key-frame, dynamic good Oba views can be collected from a scene to a scene. Although a key-frame interval can be made into 10 seconds, for example, an interval larger than this or small can also be used for it. When the interval between the key-frames which continue especially is small, a function like high-speed passing <a thing> on is realized. On the other hand, although the interval of the same size can be used for full playback of all the audios, it is only jumping video from one key-frame to the following key-frame. In this case, the enlarged display of the central key-frame is carried out. If

the key-frame of the small interval which has granularity low enough is reproduced, an expansion key-frame can be shown dynamically and high-speed passing <a thing> on (or backward feed) mode can be performed. In this example, if the material of the key-frame of the next which shows a sailing boat is reached, the line of a pars basilaris ossis occipitalis will shift only one position to the left, left end "sun" will disappear, and a new key-frame will appear from a right end. The map especially of such a display is carried out on a presentation from a background storage, and it can be performed by a frame rate quicker than standard video.

Drawing 6 shows the presentation of the subtitle within the general-purpose format attached and stated to drawing 5. . The space 50 in a center region was assigned to the actual frame, and the space 52 was taken out from the video presentation. Or it is assigned to the display of the subtitle relevant to translation in a language other than the speech text changed, other the pertinent information for hearing-impaired persons, for example, object, or the language actually used. It is not necessary to take out a subtitle only from the range relevant to seven key-frames of a screen pars basilaris ossis occipitalis. Those relevance can be extended more. Each key-frame has other associated data placed the time code 54 or on it. The two sequences 56 and 58 of a control button are assigned to application operation on left-hand side, and are assigned to the Intra program operators on right-hand side. The upper part of a screen has the title 60 of the actual video program currently displayed.

The reason constituted so that it may have a DINA MIKKURI presentation of the video cursor which moves with time in the inside of an actually active key-frame field, only by static presentation of a key-frame, he makes a video presentation dynamic as a whole, and understands deployment of an event good to a user -- making -- it is because it is insufficient. For this reason, semantics is raised as follows. in order to start that a system reduces and displays the Digital Video material including a related audio and other effects after the pause of predetermined time, the key-frame cursor "includes" becomes "alive." If it reaches during reproduction at the following key-frame, cursor "is jumped" automatically to the following key-frame shown in the user interface, and this will continue until a user starts an interaction with a system (again). Generally, the organization indicated here makes possible the browsing of information that this different from all the video strings is another. Even when playing only an audio dynamically with the jump to the arbitrary following key-frames from a key-frame, in the specific low requirements for memory, the user can acquire the good impression of presentation video.

In this point, drawing 7 is a constitutional diagram of a system action. A system waits for the input from a user in the state 100, displaying many key-frames. Such an input can include the jump between the key-frames of a large number currently displayed, the jump to another key-frame of a group, and selection of the key-frame which displays a related interval. Such arbitrary inputs drive the arrow 104 and make a new time interval start. If such an input does not have for n seconds (for example, 20 seconds), the arrow 108 will drive and the state 102 will be reached. In this state, a system performs a dynamic video cursor frame. Unless a user input is received, the arrow 110 drives, and displaying a system is continued as long as the video material which can be displayed may be obtained. However, if a user input is received, the arrow 106 will drive and a system will stop at the actual content position of a dynamic video cursor frame, or the starting position of a actual interval.

Some of key-frames which use the content of detection of a key-frame and a filtering video program influenced by the "Inta racing" effect for carrying out a browsing may be extracted from the sequence which has a high motion. As for this, a video sequence is coded by interlace coding mode like usual. When a frame consists of the two fields appointed as the composition of a perfect frame, an even line belongs to one field and an odd line belongs to the field of another side, a troublesome zigzag effect is generated. When this problem becomes much more remarkable in a small key-frame, and it becomes troublesome, and a picture is expanded in this case and a line becomes a thick block, this effect will stand out further.

First, it is necessary to detect the key-frame influenced by such an Inta racing effect. This effect is observed on the line of a picture and generates a frequency value with a high luminance change. Using this point, spatial frequency spectrum is divided into many subbands, and only a high frequency component is taken into consideration. actually, the effect which it is going to detect should show a mutual luminance value between an even line and an odd line -- it comes out, and it is, therefore this picture should have a high sampling frequency -- it comes out. It is necessary to

calculate only the coefficient which is a maximum frequency ingredient of the frequency conversion (FFT or preferably DCT) on a sequence. When a picture is influenced by a zigzag effect, it has a value with this high ingredient.

However, this effect is in sight also corresponding to the object which has a motion, especially the object which moves horizontally and has an ingredient. Therefore, total of a coefficient should not be taken into consideration. This is because total of a coefficient generates a high value in the picture which has a pattern of details and high contrast and generates an error eventually. A much more good result can be obtained by dividing a picture into some small parts and taking the maximum/zone into consideration. For example, total is no longer highly influenced by the detailed picture by adding two peak prices of each zone.

since low vertical definition is not more troublesome than a zigzag effect at the end, the easiest method of filtering this picture doubles the rise sampling of this perpendicularly only in consideration of one IRUDO. It is applicable before displaying the picture which was able to obtain the interpolation filter described with the next section.

Therefore, detection and amendment are performed as follows. The 1st step performs a twice as many rise sampling as this, in order to throw away the one field and then to restore the original size of a key-frame by removing the half of a line, even lines, or odd lines, and it performs interpolation filter processing after that. In this case, an interpolation filter performs easy linear interpolation.

In order to make it look easily from a rise sampling and interpolation TV observation distance, it is necessary to expand a key-frame to full-screen size mostly by the rise sampling which an interpolation filter follows. Generally, since a key-frame has low resolution, it is necessary to expand it for high magnification. Since a pixel will become a big block if it processes further, this means that a result will not become legible. Therefore, although it is necessary to carry out filtering of the picture, since it is necessary to carry out high speed processing so that it may have short response time while the picture of good image quality is generated and it is necessary to make it display in high resolution, it is necessary to find out a trade-off. A problem is at the point that it is necessary to expand during the flight of a picture.

That is, since the memory space needed in order to memorize to a hard disk becomes large too much, also at once, this means expansion and that filtering cannot be carried out for a picture. this -- \*\* -- It is necessary to make rise sampling and filtering processing high-speed as the result which can be permitted simultaneously is maintained and is **inside** possible Generally, The usual interpolation filter. It can be used (IEEE paper "Digital Interpolation of Discrete Image", arbitrary these ;, for example, H.C.Andrews, about digital signal processing, of C.L.Patterson).

Trans.Comput.196, v25, pp.196 -202 reference.

Other art which raises image quality can also be used. Although especially the wavelet method and the fractal method draw a high calculation burden, in vision quality, they show a remarkable result.

In practice, fractal compression art is publicly known, and can reconstruct or simulate a detail with high resolution by repeating the same decoding processing. In this case, let the picture to memorize be a fractal compression picture which generates a high compression ratio. Similarly, by using wavelet transform, a high frequency component can be predicted on a still higher scale, and high resolution images without the Japanese quince effect can be obtained.

In video transmission of the text search present of the video program based on a subtitle, A subtitle is often transmitted together with a program (when it is many, in an analogue system, it is inserted within a vertical blanking period and inserted into each elementalist beam by digital transmission). This is used to the program usually distributed in a foreign language, or is used for a hearing-impaired person. Although such information is usually superimposed on a screen, it is also recordable on a storage. If it does in this way, at the sound of a program, and the time, description of the sound for a hearing-impaired person can be used for search processing.

It is necessary to perform extraction of this kind of information in real time, recording a program. If this art is combined with a key-frame extraction routine, it is combinable with the dialog which produces a picture in a related text, i.e., the program part from which the key-frame was extracted. Thus, text browsing based on a specific keyword can be performed using the present text browsing art. The easy inquiry based on a keyword can be performed as the specific tool of application is commonly used in the "Web" search engine now.

The news program should be recorded as an example. If a keyword "France" is inserted to search the news about France, a system will look for this word in the text of a program automatically. A user is shown the key-frame about the portion of a program and the particular part of a subtitle by which the keyword was found out if a result is yes. Subsequently, the user can watch a program from a specific point. Since it is displayed on the pars basilaris ossis occipitalis of a screen as all these key-frames show drawing 5 when many key-frames are detected as a result of an inquiry, the user can analyze a related text on a window big one **at a time**. When a result is no, a keyword (France, Paris) of the same kind can be used. This system is useful also although the report which covers a specific team or a specific sport in a sports program is searched.

Many of other uses are possible, for example, it can be used for the inspection of being what a child may see in a movie, and the inspection of whether the language currently used for conversation is contained in the list of an "immoral language."

Extracting from the still picture of a screen, for example by OCR art, when - text is not obtained from video by dissociating as possible extension of such a system, extracting conversation from a program using - speech recognition technology, \*\*\*\*\*. In this case, when it operates always regardless of the service provided by the broadcast contractor and the subtitle is not provided, a system makes a system learn and can make it possible to always perform text browsing based on some keywords at least.

---

#### Drawing 1

ID=000004



#### Drawing 2

(19)日本国特許庁 (JP)

## (12)公表特許公報 (A)

(11)特許出願公表番号

特表2000-516006

(P2000-516006A)

(43)公表日 平成12年11月28日(2000.11.28)

(51)IntCl' G 06 F 17/30  
H 04 N 5/76  
5/93

識別記号

F I  
G 06 F 15/40  
H 04 N 5/76  
G 06 F 15/403  
H 04 N 5/93

テ-ヨ-ト<sup>\*</sup> (参考)  
3 7 0 D  
B  
3 8 0 F  
Z

審査請求 未請求 予備審査請求 未請求(全 24 頁)

(21)出願番号 特願平10-529208  
(86) (22)出願日 平成10年1月22日(1998.1.22)  
(85)翻訳文提出日 平成10年10月2日(1998.10.2)  
(86)国際出願番号 PCT/IB98/00091  
(87)国際公開番号 WO98/34182  
(87)国際公開日 平成10年8月6日(1998.8.6)  
(31)優先権主張番号 97200286.9  
(32)優先日 平成9年2月3日(1997.2.3)  
(33)優先権主張国 ヨーロッパ特許庁 (EP)  
(31)優先権主張番号 97201802.2  
(32)優先日 平成9年6月16日(1997.6.16)  
(33)優先権主張国 ヨーロッパ特許庁 (EP)

(71)出願人 コニンクレッカ フィリップス エレクトロニクス エヌ ヴィ  
オランダ国 5621 ベー-アーア アインドル  
フェン フルーネヴァウツウェッハ 1  
(72)発明者 セッカレリ マルコ ポーロ  
オランダ国 5656 アーアー アインドル  
フェン ブロフ ホルストラーン 6  
(74)代理人 弁理士 杉村 明秀 (45名)

最終頁に続く

(54)【説明の名称】複数のキーフレームを並列に表示することによりビデオ内容を案内する方法及び装置

## (57)【要約】

1列以上の複数のキーフレームを並列に表示することによりビデオ内容を案内するとともに、表示されたキーフレームの選択のアクセスにより前記ビデオ内容への実際のアクセスをこのようにアクセスしたキーフレームのマッピングを表しながら制御し得る方法である。この方法は、単一のユーザインターフェース構成において、キーフレームをスクリーン上に時間的に整列した順序に配列する第1動作モードと、表示される連続キーフレーム間に多数の選択可能なグラニュラリティを有するキーフレームを配列する第2動作モードとの間で選択を行うことができる特徴を有する。

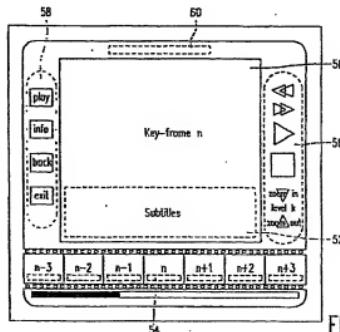


FIG. 6

**【特許請求の範囲】**

1. 1列以上の複数のキーフレームを並列に表示することによりビデオ内容を案内するとともに、表示されたキーフレームの選択的アクセスにより前記ビデオ内容への実際のアクセスをこのようにアクセスしたキーフレームのマッピングを表しながら制御し得る方法において、

单一のユーザインターフェース編成内において、キーフレームをスクリーン上に時間的に整列した順序に配列する第1動作モードと、表示される連続キーフレーム間に不均一の選択可能なグラニュラリティを有するキーフレームを配列する第2動作モードとの間で選択を行うことができる特徴とするビデオ内容案内方法。

2. 前記時間的に整列した順番において時間的に中心に位置するキーフレームに関連するオーディオインターバルを順次に再生することを特徴とする請求項1記載の方法。

3. 順次のオーディオインターバルが離散分離したキーフレームのシーケンスに対しほぼ連続したオーディオ表現を構成することを特徴とする請求項2記載の方法。

4. 第2動作モードにおいて、実際にアクセスされたキーフレームに関連するオーディオインターバルを再生することを特徴とする請求項1記載の方法。

5. 現在選択されているキーフレームをハイライト表示すると同時に、このキーフレームを他のキーフレームに対し多数倍の大きさのフォーマットに拡大し、更に、有害なビデオインクレーシング効果を検出し、このような効果が検出される場合には、このような効果を垂直デシメーションにより軽減させることを特徴とする請求項1記載の方法。

6. 現在選択されているキーフレームをハイライト表示すると同時に、このキーフレームを他のキーフレームに対し多数倍の大きさのフォーマットに拡大し、更に、画像を表示する前に画像にアップサンプリングフィルタ処理を施すことを特徴とする請求項1記載の方法。

7. 関連するキーフレーム又はキーフレームのシーケンスに対し抽出されたサブ

タイトル又は他の関連情報を表示することを特徴とする請求項1記載の方法。

8. 請求項1に記載された方法を実行するよう構成された装置。

## 【発明の詳細な説明】

複数のキーフレームを並列に表示することによりビデオ内容を案内する方法及び装置

### 発明の背景

本発明は、1列以上の複数のキーフレームを並列に表示することによりビデオ内容を案内するとともに、表示されたキーフレームの選択的アクセスにより前記ビデオ内容への実際のアクセスをこのようにアクセスしたキーフレームのマッピングを表しながら制御し得る方法に関するものである。後の選択的再生のために記録されるビデオプレゼンテーションの代表的部分をキーフレームとして使用することが提案されている。連続ビデオストリームとはビデオが“オン”的まであることを意味し、このストリームにはアニメーション、一連の静止画、又は画像のインタラクティブシーケンスを含めることができる。そのキャラクタは種々のもの、例えば映画、ニュース、又はショッピングリストとすることができる。この技術はS. W. Smoliar及びH. J. Zhangの論文”Content-Based Video Indexing and Retrieval”、IEEE Multimedia, Summer 1994, pp.62-72に開示されている。

キーフレームはユーザ宅においてビデオ材料の受信時にこの材料から導出アルゴリズムにより取り出すことができ、或いは例えばビデオプロバイダにより、各ビデオショットがキーフレームで開始するようにキーフレームにラベルを付けることができる。第3の方法は、これらのフレームを標準ビデオ速度に関連する均一の時間インターバルで互いに連続させる。本発明は、キーフレームを使用してビデオプレゼンテーションのダイナミックなオーバビューを有用なファシリティとともにユーザに与え、ビデオ材料へのアクセス、次の表示の選択又は選択解除、又は編集を一層容にする必要があるという認識に基づくものである。

ビデオ画像のデジタル及び圧縮符号化に対する現在のプロジェクトに対する特定の問題は、大容量媒体へのその記憶は一般に即時アクセスを許容せず、特に単位記憶サイズ当たりのフレーム数で表せるリニア記憶密度が不均一である点に

ある。テープのような高容量主記憶媒体に、小容量であるが高いアクセス速度を

有する補助記憶媒体を付加することが提案されている。この場合には、高速順送り及び高速逆送りのようなトリックモードの実行、並びに次の表示ためにビデオ材料を要約形、変更形、又は再配列形に編集する処理がユーザインターフェースの点及び記憶技術の点の双方から見てかなりの難点を生ずる。

### 発明の概要

従って、本発明の目的は、特に、ユーザインターフェース編成に一層高いフレキシビリティを導入するとともに、ユーザに記憶編成並びにビデオ材料の一層自然なフィーリングを与えることにある。この目的のために、本発明の特徴は、単一のユーザインターフェース編成内において、キーフレームをスクリーン上に時間的に整列した順序に配列する第1動作モードと、表示される連続キーフレーム間に複数の不均一の選択可能なグラニュラリティを有するキーフレームを配列する第2動作モードとの間で選択を行うことができる。キーフレームを時間的に整列した順序で提示すると、例えばこれらのフレームが標準ビデオ速度に関連する均一の時間インターバルで互いに連続する場合には、高速順送り及び高速逆送りを簡単に実行することができる。更に、フレーム間における可変グラニュラリティを有する階層レベルの簡単容易な変化により簡単容易なアクセス及び編集を行うことができる。キーフレーム又は少なくともその一部分を映画ショットの開始部から取り出す場合、又はオリジナルフィルムエディタにより発生された他の関連するイベントから取り出す場合にも同じことがいえる。このようにして、クラスタリング処理を自動的に実行することができる。

本発明方法においては、現在選択されているキーフレームを他のキーフレームに対し多数倍のフォーマットで拡大することによりこのフレームをハイライト表示するとともに、有害なビデオインターレス効果を検出し、もしそうであれば、垂直デシメーションによりこの効果を低減させ、且つ又表示前に画像にアップサンプリングフィルタ処理を施す。比較的小さいキーフレームのビデオ歪みは許容し得ることが経験上解っているが、特定のキーフレームを拡大する場合には、画像改善のために追加の手段を講ずる必要がある。本発明者は、このアップグレーディングは、標準状態下的画質に必ずしも達しないが、画質の快適で有益な向上

をもたらすことを確かめた。

本発明は上述した方法を実施するよう構成した装置にも関するものである。本発明の他の特徴は従属請求項に記載されている。

#### 図面の簡単な説明

本発明のこれらの特徴及び他の特徴及び利点を図面を参照して好適実施例について以下に詳細に説明する。図面において、

図1はTVレコーダ組合せ装置のブロック図であり、

図2はビデオ記録の代表的構造を示し、

図3はスクローリングモザイクユーザインターフェースの設計例を示し、

図4はスクローリングリストユーザインターフェースの設計例を示し、

図5はもっと広範なグラフィカルユーザインターフェースの設計例を示し、

図6はサブタイトルの表示を示し、

図7はシステム動作の状態図を示す。

#### 好適実施例の詳細な説明

特に、通常のカスタマ及びプライベートな家庭の使用に關係するが、このようないくつかの使用に關係するものに限定されるものではない種々の利点は次のとおりである。

- ・キーフレームは代表的なTV観察距離に位置するユーザがこれらを互いに識別し得るように提示させなければならない。
- ・同時に提示するキーフレームの数はユーザにデジタルビデオ材料のコンテンツの有意部分のオーバビュを与えるのに十分とすべきである。
- ・キーフレームは、例えばアスペクト比を保持することにより無歪みの形で表示すべきである。
- ・TVセットのリモートコントロール装置自体がユーザコントロール装置として動作するようにするのが好ましい。
- ・フィードバック情報は代表的観察距離から認知できなければならない。
- ・“ドラッグ”及び“ドロップ”的なコンピュータコンセプトは一般に不要にする。・種々の機能は連続的ではなく時折使用されるだけとなるように実現しなければならない。

・ユーザインターフェースはビデオプレゼンテーションの良く知られたリニアモデルを表すべきである。

#### 特定の実施例の説明

図1は本発明によるTVレコーダ組合せ装置を示す構成図である。要素20はTVセットディスプレイ及び関連する即時制御兼電源部を示す。要素22はアンテナ、又はケーブル配信のような他のタイプの信号分配エンティティを有する接続部を示す。この要素には受信信号からデジタルビデオ情報又はデジタル信号部分を取り出すエンティティも適宜含めることができる。要素34は図1の種々のサブシステム間のビデオストリーム及び関連する情報の経路指定部を示す。経路指定は制御ポックス28によってライン35を介して制御信号により制御することができる。ライン35は単一の双方向相互接続ラインとして示されているが、実際には任意の数の单方向又は双方向ラインからなるものとすることができる。制御ポックス28はディスプレイ20からライン30を経て検出信号を受信するとともに他のサブシステム38、40から検出信号を受信し、これらのサブシステムを制御する。ブロック38はマルチギガバイト領域の極めて高い記憶容量を有するリニアテープレコーダである。ブロック40は高い記憶容量を有する磁気ディスクレコーダであり、このレコーダはレコーダ38の記憶容量の一部分を有するにすぎないが、このレコーダのアクセス速度はクロストラックジャンプアクセスによりレコーダ38よりはるかに速い。ブロック38とブロック40が一緒にコンピュータメモリキャッシュシステムに類似の2レベル記憶編成を構成し、ビデオプレゼンテーションのすべてのアイテムを少なくとも1度記憶する。要素24はディスプレイ装置20とワイヤレスバス26を介して通信するとともにサブシステム28及び他のサブシステム38及び40と間接的に通信するリモートコントロール装置を示す。

図2はビデオプレゼンテーションの代表的な構成を示す。ビデオ内容を機能的にするために、バー60はビデオ自体を、フレームの形で、又はMPEG符号のような圧縮ビデオ内容のストリングとして含む。情報はバーに沿ってビデオ時間の進行とともに記憶されるが、実際の記憶要件は再生時間に亘って均一にする必

要はない。点在するキーフレームが68のような黒垂直ストライプで示されている。各キーフレームは次のキーフレームまでのインターバル内の全ビデオを表す

もの、又はこれを代表するものとして使用される。キーフレームは、ビデオバイダがラベルを付ける又は”テーブル・オブ・コンテンツ”(TOC)を挿入することにより各新ショットの第1フレームとして選択することができる。或いは又、レシーバが、あるアルゴリズムにより、ビデオコンテンツが一つのフレームから次のフレームへ急激に変化することを検出することができる。本発明はこのアルゴリズムが正しいものと仮定する。図に示すように、キーフレームの分布は不均一にすることができる。他のメカニズムは、連続するキーフレームを規定のインターバル、例えば2-3秒ごとに互いに連続させることにある。本例では、表示部62にキーフレームのみを示す。更に、キーフレームは幾分階層構造に編成され、表示部64は限定された一組の高レベルのキーフレームのみを示す。この階層構造は多レベルにすることができ、表示部66はビデオプレゼンテーション60の全体に対する単一のキーフレームのみと関連するものとすることができる。キーフレームの種々のレベルは上述した編成と異なる編成内で定めることができ、並べることもできる。

図1において記憶マッピングを、ビデオプレゼンテーションの本体をテープレコーダ38に記憶するとともに、少なくともキーフレームをできれば関連するキーフレームの直後の短いビデオ及び/又はオーディオインターバルと一緒にディスクレコーダ40で再生するように行うことができる。このようなインターバルの長さはリニアテープレコーダ38の待ち時間に対応させて、リアルタイムアクセスを達成することができる。ビデオプレゼンテーションは映画のように本質的にリニアなものとし得る。他の使用では、所定の記憶インターバルにアニメーション、静止画、又はコンシューマに使用される他の画像を含めることができる。あるキーフレームはこれを抑圧することができる。これは、当該キーフレームの前の時間インターバルをその後の時間インターバルと有効に結合する。リセット機能によりインターバルを再び分離することができる。また、種々のクラスのキーフレーム、例えば一定の時間インターバルで互いに分離されたクラスのキーフ

レームを抑圧することができる。種々の異なるクラスのキーフレーム、例えば受信時にローカルアルゴリズムにより発生されるキーフレームに対しプロバイダにより導入されるキーフレームを1つのプレゼンテーションに使用することができ

る。

図3はスクローリングモザイクユーザインターフェースの設計例を示す。どのスクリーンも左上から出発して右下まで20個のキーフレームを提示する。各キーフレームは図示のキーフレームの総合順位の番号を有している。実際には、キーフレーム144が矩形の制御カーソルによりハイライト表示されている。ユーザはリモートコントロールを駆動し、カーソル装置のナビゲーション制御部によってカーソルを表示されたキーフレーム上及び上部及び底部に表示されたバーに表示されたボタン上を自由に移動させることができる。ユーザが制御カーソルを左上コーナにおいて左に移動させると、表示が20キーフレームだけ後ろにジャンプする。右下コーナにおいて右に移動させると、表示が20フレームだけ前にジャンプする。スクリーンの上部バーをアクセスすることにより5つの等しい長さの部分に分割されたプレゼンテーションの他の部分のアクセスを制御することができ、黒い水平バーが全プレゼンテーションのうちの、ここに表示された20個のキーフレームによりカバーされる全時間を示す。

最初に特定のキーフレームを選択し、次いで底部ボタンの1つを選択することにより他の機能を開始させることができる。"view program"(プログラムビュー)ボタンはカーソルでアクセスしたキーフレームでの開始を制御する。"View segment"(セグメントビュー)ボタンはビュープログラムと同一であるが、次のキーフレームで終了する単一セグメントのみを再生する。"view from x to y"(xからyまでビュー)ボタンはカーソルでアクセスした2つのキーフレームのうちの時間的に早い方のフレームでの開始と、時間的に遅い方のフレームでの停止を制御する。他のモードをキーフレーム選択機能とともに実現することができる。例えば、ユーザが特定のインターバルの発生をチェックできるようにするFast-forward(高速順送り)又はslow-forward(低速順送り)、又は所定のビデオ効果を達成するfast/slow reverse(高速/低速逆送り)を達成することができる。表示中に

、特定のキーフレームに関連する瞬時の通過時に、このキーフレームがアクティブになり、次のキーフレームと関連する瞬時に到達するまでビデオストリームを有効に表示する。次のキーフレームの到達時に、このフレームがアクティブフレームになる。このような機能によって、ユーザはビデオレコーダをインタ

ーバル表示シーケンスに対し、例えばコマーシャルのような所定のセグメントを削除することにより、低速順送りによって所定のディテールに注意が向くようにストレートフォワードにプログラムすることが可能になる。表示中に、図示していない制御ボタンによりオーディオをアクティブにしたり、抑圧したりすることができる。或いは又、オーディオを制御し続けるが、ビデオカーソルはディスクリートにして、適切なハイライト指示によりインターバルからインターバルにステップするのみとすることもできる。

図4はスクローリングリストユーザインターフェースの代表的な設計例を示す。本例では、全スクリーンはその底部に表示された5つのキーフレームを有し、キーフレーム145がそのエッジに沿って移動する矩形の制御カーソルによりハイライト表示されている。キーフレーム145は大きな倍率でバックグラウンド内にも表示される。この制御インターフェースは図3のものと同一であるが、ボタンの位置が相違している。拡大キーフレームはマルチキーフレームバー内に抑圧することもできる。

図5はもっと広範なグラフィカルユーザインターフェースを示す。第1に、左右のコラムは、play(再生)、stop(停止)、select(選択)、cut(切り取り)、paste(張り付け)、fast reverse(高速逆送り)、zoom+(ズーム+)、zoom-(ズーム-)、fast forward(高速順送り)のための制御ボタンのコラムである。底部の行はほとんど相関を有しないそれぞれ異なるシーン又はショットに関する9つのキーフレームのシーケンスを有する。キーフレームの階層構造内のステップ移動により、シーンからシーンへダイナミックな良好なオーバビューを集めることができる。キーフレーム間隔は、例えば10秒にすることができるが、これより大きい又は小さい間隔を使用することもできる。特に連続するキーフレーム間の間隔が小さい場合には、高速順送りのような機能が実現される。他方

、同一の大きさの間隔を全オーディオのフル再生のために使用することができるが、ビデオは1つのキーフレームから次のキーフレームへジャンプするのみである。この場合、更に中心キーフレームが拡大表示される。十分に低いグラニュラリティを有する小間隔のキーフレームを再生すると、拡大キーフレームをダイナミックに提示することができ、高速順送り（又は逆送り）モードを実行することができます。

きる。本例では帆船を示す次のキーフレームの材料に到達すると、底部の行が1位置だけ左にシフトし、左端の“太陽”が消え、新しいキーフレームが右端から現れる。このような表示は特に背景記憶媒体からプレゼンテーション上にマップされ、標準ビデオより速いフレームレートで行うことができる。

図6は図5につき述べた汎用フォーマット内のサブタイトルのプレゼンテーションを示す。中央領域内のスペース50が実際のフレームに割り当てられ、スペース52がビデオプレゼンテーションから取り出された、又は他の関連情報、例えば聴覚障害者用に変換されたスピーチテキスト又は実際に使用されている言語とは別の言語への翻訳に関連するサブタイトルの表示に割り当てられている。サブタイトルはスクリーン底部の7つのキーフレームに関連する範囲からのみ取り出す必要はない。それらの関連性はもっと広げることができる。更に、各キーフレームはタイムコード54又はその上に置かれる他の関連データを有する。制御ボタンの2つの列56、58が左側にアプリケーション操作用に、右側にインタラクティブプログラムオペレータ用に割り当てられている。スクリーンの上部は表示されている実際のビデオ番組のタイトル60を有する。

実際にアクティブなキーフレーム領域内を時間とともに移動するビデオカーソルのダイナミクリプレゼンテーションを有するように構成する理由は、キーフレームのスタティック提示だけではビデオプレゼンテーションを全体としてダイナミックにしてユーザにイベントの展開を良好に理解せしめるには不十分であるからである。このためにセマンティクスを次のように向上させる。システムは、所定の時間の休止後に、関連するオーディオ及び他の効果を含むデジタルビデオ材料を縮小して表示することを開始するため、カーソルが“内包する”キーフ

ームが“アライブ”になる。再生中に、次のキーフレームに到達すると、カーソルはユーザインターフェース内に提示されている次のキーフレームに自動的に“ジャンプ”し、これはユーザがシステムとのインタラクションを(再び)開始するまで続く。一般に、ここに記載する構成は全ビデオストリングと異なるこれとは別の情報の走査検索を可能にする。任意のキーフレームから次のキーフレームへのジャンプとともにオーディオのみをダイナミックに再生する場合でも、ユーザは特定の低い記憶要件において提示ビデオの良好な印象を得ることができる。

この点において、図7はシステム動作の状態図である。状態100において、システムは多数のキーフレームを表示しながらユーザからの入力を待つ。このような入力は表示されている多数のキーフレーム間のジャンプ、別の組のキーフレームへのジャンプ、関連するインターパルを表示するキーフレームの選択を含むことができる。任意のこのような入力は矢104を駆動し、新しい時間インターパルを開始させる。このような入力がn秒(例えば20秒)間ないと、矢108が駆動され、状態102に到達する。この状態において、システムはダイナミックなビデオカーソルフレームを実行する。ユーザ入力が受信されない限り、矢110が駆動され、システムは表示可能なビデオ材料を入手し得る限り表示しつづける。しかしユーザ入力が受信されると、矢106が駆動され、システムはダイナミックなビデオカーソルフレームの実際のコンテンツの位置、又は実際のインターパルの開始位置に留まる。

#### “インターシング”効果により影響されたキーフレームの検出及びフィルタリング

ビデオプログラムのコンテンツを走査検索するのに使用するキーフレームのいくつかは高い動きを有するシーケンスから抽出されたものであることがある。これは、ビデオシーケンスが通常の如くインターレース符号化モードで符号化され、フレームが完全なフレームの構成に起用する2つのフィールドからなり、偶数ラインが一方のフィールドに属し、奇数ラインが他方のフィールドに属する場合には、わずらわしいジグザグ効果を発生する。この問題は小さいキーフレームにおいて一層顕著になり、わずらわしくなり、この場合には画像が拡大され、ライン

が太いブロックになると、この効果が一層目につくことになる。

最初に、このようなインターレーシング効果により影響されたキーフレームを検出する必要がある。この効果は画像の行上に観察され、輝度変化が高い周波数値を発生する。この点を利用して、空間周波数スペクトルを多数のサブバンドに分割し、高い周波数成分のみを考慮する。実際には、検出しようとする効果は偶数ラインと奇数ラインとの間で交互の輝度値を示すはずであり、従ってこの画像は高いサンプリング周波数を有するはずである。列上の周波数変換（FFT又は好ましくはDCT）の最高周波数成分である係数のみを計算する必要がある。画像

がジグザグ効果により影響されると、この成分が高い値を有する。

しかし、この効果は動きを有する物体、特に水平方向に動き成分を有する物体にも対応して見える。従って、係数の総和は考慮すべきでない。その理由は、係数の総和は詳細且つ高コントラストのパターンを有する画像において高い値を発生し、最終的に誤りを発生するからである。画像をいくつかの小部分に分割し、最大値／区域を考慮することにより一層良好な結果を得ることができる。例えば、各区域の2つの最高値を加算することにより総和が高度に詳細な画像に影響されなくなる。

最後に、低い垂直解像度はジグザグ効果よりわざらわしくないので、この画像をフィルタリングする最も簡単な方法は1つのイールドのみを考慮し、これを垂直方向に2倍にアップサンプリングするものである。次のセクションで述べる補間フィルタを得られた画像を表示する前に適用することができる。

従って、検出及び補正是次のように行われる。第1ステップは行の半分、偶数行又は奇数行、を除去することにより1つのフィールドを捨て、次にキーフレームの元のサイズを復元するために2倍のアップサンプリングを行い、その後に補間フィルタ処理を行う。この場合には、補間フィルタは簡単なリニア補間を行う。

#### アップサンプリング及び補間

TV観察距離から容易に見えるようにするために、キーフレームは補間フィルタが後続するアップサンプリングによりほぼフルスクリーンサイズに拡大する必

要がある。一般にキーフレームは低い解像度を有するので、高い倍率で拡大する必要がある。これは、更に処理すると、画素が大きなブロックになるので、結果は見やすいものにならないことを意味する。従って、画像をフィルタ処理する必要があるが、良好な画質の画像を発生させ、高い解像度で表示させる必要がある一方、短い応答時間を持つように高速処理する必要があるため、トレードオフを見出す必要がある。問題は、拡大を画像の飛行中に行う必要がある点にある。即ち、このことは、ハードディスクに記憶するためには必要とされる記憶スペースが大きくなりすぎるため画像を一度も拡大及びフィルタ処理することができないことを意味する。これがため、アップサンプリング及びフィルタリング処理は、同時に許容し得る結果を維持しながらできるだけ高速にする必要がある。一般に、

通常の補間フィルタを使用することができる（デジタル信号処理に関する任意の本：例えばH.C.Andrews, C.L.Pattersonの論文“Digital Interpolation of Discrete Image”, IEEE Trans. Comput. 196, v25, pp.196-202参照）。

画質向上させる他の技術を使用することもできる。特に、ウェーブレット法及びフラクタル法は高い計算負担を専らが視覚品質において顕著な結果を示す。実際に、フラクタル圧縮技術は公知であり、同一の復号処理を反復することによりディテールを高解像度で再構成する又はシミュレートすることができる。この場合には、記憶する画像を高い圧縮率を発生するフラクタル圧縮画像とする。同様に、ウェーブレット変換を使用することにより、一層高いスケールで高周波数成分を予測して、ボケ効果のない高解像度画像を得ることができる。

#### サブタイトルに基づくビデオプログラムのテキストサーチ

現在のビデオ伝送においては、番組と一緒にサブタイトルがしばしば伝送される（多くの場合、アナログシステムでは垂直プランギング期間内に挿入され、デジタル伝送では個々のエレメンタリストリーム内に挿入される）。これは通常外国語で配信される番組に対し使用され、或いは聴覚障害者のために使用される。このような情報は通常スクリーン上に重畳されるが、記憶媒体に記録することもできる。このようにすると、番組の音声及び時には聴覚障害者のための音声の

記述をサーチ処理に利用することができる。

この種の情報の抽出はプログラムを記録しながらリアルタイムで行う必要がある。この技術をキーフレーム抽出ルーチンに結合すると、画像を関連するテキスト、即ちキーフレームが抽出された番組部分に生ずるダイアログに結合することができる。このように、現在のテキスト検索技術を用いて特定のキーワードに基づくテキスト検索を実行することができる。アプリケーションの特定のツールが、現在“Web”サーチエンジンにおいて慣用されているように、キーワードに基づく簡単な問合せを実行することができる。

一例として、ニュース番組が記録されたものとする。フランスに関するニュースを検索したい場合には、キーワード“フランス”を挿入すると、システムは自動的に番組のテキスト内のこのワードを探す。結果がイエスなら、ユーザにキーワードが見つけ出された番組の部分及びサブタイトルの特定部分に関するキーフ

レームが提示される。次いでユーザは番組を特定の点から見ることができる。問合せの結果として多数のキーフレームが検出される場合には、これらのキーフレームのすべてが図5に示すようにスクリーンの底部に表示されるため、ユーザは関連するテキストを1つづつ大きなウィンドウ上で分析することができる。結果がノーの場合には、同種のキーワード（フランス、パリ）を使用することができる。このシステムはスポーツ番組において特定のチーム又は特定のスポーツをカバーするレポートを検索するにも有用である。

多くの他の用途が可能であり、例えば映画が子供が見てもよいものか否かの検査、会話に使用されている言葉が“不道德言語”的リストに含まれているか否かの検査に使用することができる。

このようなシステムの可能な拡張として、

- ・テキストがビデオから分離して得られない場合には、スクリーンの静止画から、例えばOCR技術により抽出すること、
- ・音声認識技術を用いて番組から会話を抽出すること、

がある。この場合には、システムは放送業者により提供されるサービスと常に無関係に動作し、サブタイトルが設けられていない場合には、システムを学習させ

て少なくともいくつかのキーワードに基づいてテキスト検索を常に実行し得るようになることができる。

【図1】

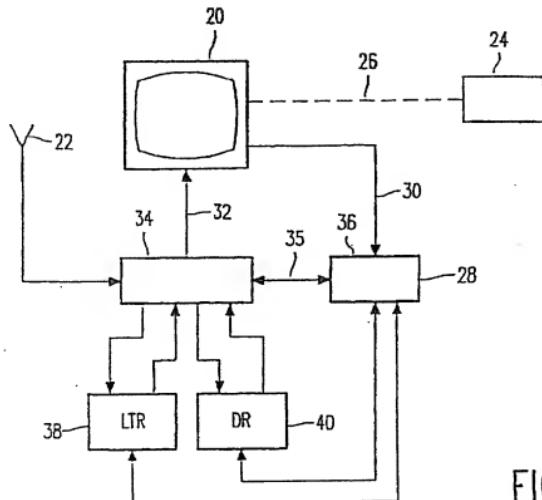


FIG.1

【図2】

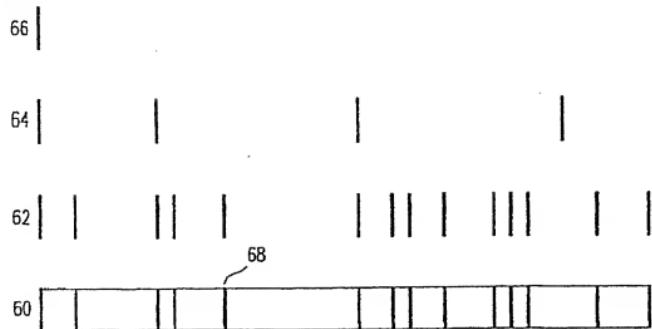


FIG.2

【図3】

	0%	20%	40%	50%	60%	80%	100%
key frame	138	139	140	141	142		
143	144	145	146	147			
148	149	150	151	152			
153	154	155	156	157			
view program	view segment	view from x to y	...	...			

FIG.3

【図4】

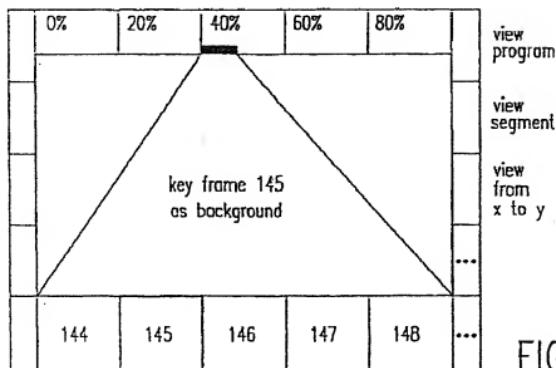


FIG.4

【図5】

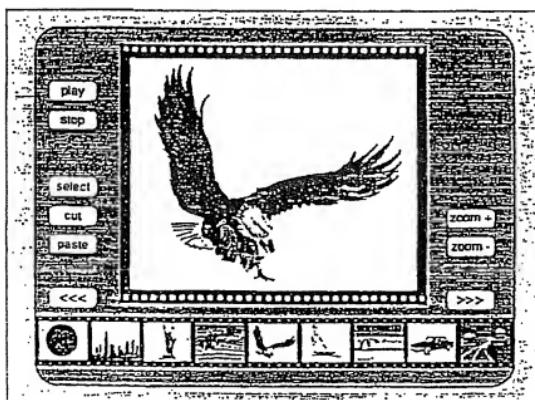


FIG.5

【図6】

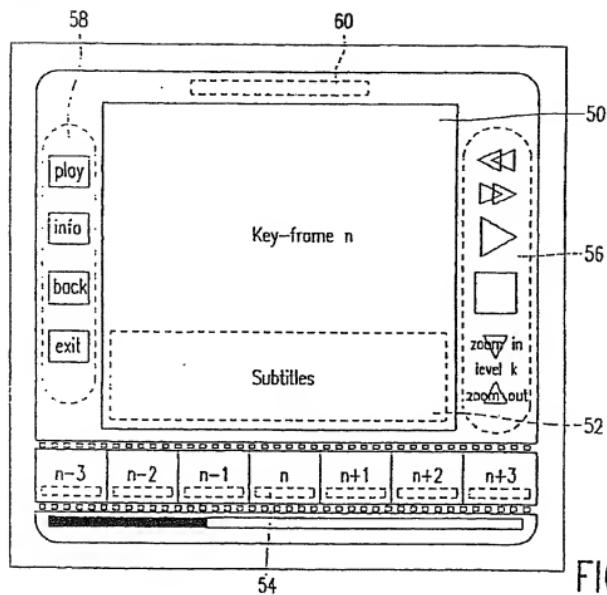


FIG. 6

【図7】

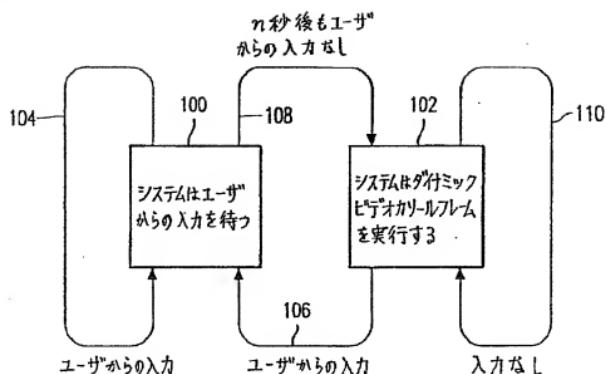


FIG.7

## 【国際調査報告】

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/IB 98/00091

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
IPC6: G06F 17/30 According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols)		
IPC6: G06F Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched <b>SE, DK, FI, NO classes as above</b>		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
<b>WPIL, EDOC, INSPEC</b>		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Journal of Visual Languages and Computing, Volume 1, 1990, Y. Tononura et al., "Content Oriented Visual Interface Using Video Icons for Visual Database Systems", page 183 - page 198, see the whole document	1-8
X	IEEE Transactions on Consumer Electronics, Volume 41, No 3, August 1995, H-Y Chen et al., "A multi-layer video browsing system", page 842, see the whole document	1-8
X	IEEE Multimedia, Summer 1994, pages 62-72, S. W. Smoliar et al: "Content-Based Video Indexing and Retrieval", see the whole document	1-8
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.		<input checked="" type="checkbox"/> See patent family annex.
<p>* Special categories of cited documents</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"B" other document but published on or after the international filing date</p> <p>"C" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"D" document referring to an oral disclosure, use, exhibition or other special exhibition</p> <p>"E" document published prior to the international filing date but later than the priority date claimed</p>		
<p>"T" later document published after the international filing date or priority date and not in conflict with the application, has cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>		
Date of the actual completion of the international search	Date of mailing of the international search report	
<b>10 August 1998</b>	<b>14-08-1998</b>	
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer <b>Bo Gustavsson</b> Telephone No. +46 8 782 25 DU	

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB 98/00091

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Proceedings of the SPIE, Volume 2417, 1995, H.J. Zhang et al., "Content-Based Video Browsing Tools", page 389 - page 398, see the whole document --	1-8
A	EP 0739129 A2 (EASTMAN KODAK COMPANY), 23 October 1996 (23.10.96), page 2, line 21 - line 33 -----	5

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

27/07/98

International application No.

PCT/IB 98/00091

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0739129 A2	23/10/96	JP 8307820 A US 5579054 A	22/11/96 26/11/96

---

フロントページの焼き

(81)指定国 EP(AT, BE, CH, DE,  
DK, ES, FI, FR, GB, GR, IE, IT, L  
U, MC, NL, PT, SE), JP, KR